

UNCLASSIFIED

AD 414859

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

9895
AFOSR

62-45

This research was supported by the
Biological Sciences Division, AFOSR,
SRLA

under Contract/Grant **AF-AFOSR 62-117**

~~(AF-AFOSR 37063)~~

**FINAL
SUMMARY REPORT**

Air Force Grant No. AF-AFOSR 62-117

Submitted by the Committee on Mathematical Biology
The University of Chicago

**DDC
RECEIVED
AUG 31 1962
REGISTERED
TISA E**

SUMMARY OF COMPLETED WORK - March 1, 1962 to March 1, 1963

The following summary includes work which has been completed or brought to a fairly complete form. Various parts of the following are being prepared for publication. Some proposals for future work are also included.

SOME APPLICATIONS OF NEURAL NET THEORY OF LEARNING

A number of problems involved in learning have been considered from the point of view of neural net theory. In this theory it is assumed that the effect of learning is only to increase an excitation at some center, this excitation increasing in some very simple way with a number of reinforcements. In most of the examples it will be assumed that this increase is linear. We shall generally ignore the effects which arise due to the fact that there are stochastic aspects involved. It should be emphasized here that in the stochastic models of learning it is generally assumed that the probabilities change as a result of learning (through some sampling mechanism perhaps) in a linear manner. On the other hand, in the present model it is assumed that the learning changes an excitation or, what is equivalent, results in a change in depolarization of neural membranes. One could, of course, include here a change in thresholds since one cannot tell the difference between an increased excitation or a decreased

414859

**CATALOGED BY DDC
AS AD No. 414859**

threshold. If there is a single stimulus involved, then one simply has a threshold, a change in the excitation and a random fluctuation from which to determine the change in probability of the response. On the other hand if there are two stimuli between which a choice must be made, then the threshold, itself, does not enter, but only a change in excitation enters, together with the fluctuations. There then results a discrimination between the excitations corresponding to two stimuli, either of which may have been altered by conditioning.

On the basis of such a simple system it then seems reasonable to attempt to account for some of the classical learning data to see whether or not the model can account for the data in a simple way. When the situation involved is that of avoidance, then it will be assumed that the conditioning takes place between the conditioned stimulus and whatever response the unconditioned stimulus first initiates. That is, for example, if the unconditioned stimulus initiates a fear or flight reaction, then it is assumed that conditioning takes place between the conditioned stimulus and this flight reaction. It will be assumed that this flight reaction affects primarily a center which initiates random movements and, therefore, contributes to the random fluctuations or perhaps lowers thresholds to responses in general. Thus, in avoidance learning there are likely to be two processes, 1) a conditioning to excitement (or perhaps, withdrawal) and 2) a conditioning to the correct response. These processes may, in principle, be separated experimentally. Thus it can be seen that a parameter involved in simple discrimination enters into the expressions for performance.

1. Learning under Partial Reinforcement.

We consider the situation in which there is a choice of two responses, each of which is rewarded with some fixed probability, π_c and π_w being these probabilities. Using the simplest model, assuming linearity and neglecting any possibility of extinction, we have for the mean value of ε_c and ε_w after n trials, $\varepsilon_c = \varepsilon_{oc} + b\pi_c c$ and $\varepsilon_w = \varepsilon_{ow} + b\pi_w w$, where c is the number of "correct" responses, w the number of "wrong" responses and $c + w = n$. If we also assume $\varepsilon_{oc} = \varepsilon_{ow}$ and π_c sufficiently greater than π_w so that the probability of the correct response P_c satisfies $P_c > \frac{1}{2}$, then we have

$$P_c(n) = 1 - \frac{1}{2} e^{-kb(\pi_c c - \pi_w w)} \quad (1.1)$$

Now using the continuous approximation $\frac{dc}{dn} = P_c$, and $w = n - c$, we can obtain the expression for P_c as a function of n ,

$$P_c = \frac{\pi_c - \pi_w + \pi_w e^{-kb\pi_c n}}{\pi_c - \pi_w + (\pi_c + \pi_w) e^{-kb\pi_c n}} \quad (1.2)$$

Note that $P_c(0) = \frac{1}{2}$ and $P_c(\infty) = 1$, and $\left(\frac{dP_c}{dn}\right)_0 = \frac{1}{4}b(\pi_c - \pi_w)$.

Consider three special cases: a) $\pi_c = 1$, $\pi_w = 0$; b) $\pi_c = 0.5$, $\pi_w = 0$; and c) $\pi_c = .75$, $\pi_w = .25$. We see from (1.2) that by plotting P_c against $(\pi_c - \pi_w)n$ all three curves should start out together, cases (a) and (b) then resulting in the same curves. The

curve from case (c) should, according to this model, soon give somewhat higher values of P_C .

Stanley⁽¹⁾ gave data on T-maze experiments with rats using three reward schedule. Seven rats were used in each of the three groups and eight trials run each day. These data (adjusted as explained below) are plotted in Fig. 1 as P_C , percent correct turns during each day, against $(\pi_C - \pi_W)n$. It can be seen that the data agree with the prediction that the initial slopes should be the same for the three cases, and also that groups (a) and (b) give essentially the same curve.

Unfortunately Stanley ran each rat only until the animal went two days, not necessarily consecutive, without errors. This means that the recorded values of P_C , for about the last half of each group, are in error since they do not include the rats that had presumably reached 100% correct response. We have adjusted his data on the assumption that the rats that were withdrawn would have continued to perform without error. Using this adjusted data, it can be seen from the figure, that the prediction of higher values for case (c) is also confirmed.

2. Avoidance Learning.

A neural net model for avoidance learning is considered. This is intended to represent the experiments performed by Brady and Marascio (Psychological Record 12, 361-365, 1962) in which a rat is placed in a box containing food and water and left in complete darkness for an extended period. At one end of the box electrodes are arranged so as to

⁽¹⁾ Stanley, J.C., Jr., Ed. D. Thesis, Harvard University, 1950. Quoted in Bush and Mosteller, Stochastic Models for Learning, p. 252.

make a record of contacts. In a control run, only the record of contacts is made, but in an experimental run the contacts are recorded and the animal simultaneously receives a shock.

For the control situation, it is assumed that there is an excitation $\epsilon(t)$, due to the tendency to explore a new environment, which decreases linearly with time during a period $0 \leq t \leq 1/\theta$, to a constant base level, ϵ_0 . The response (entering the other end of the box) occurs when $\epsilon(t)$ plus a random excitation, ξ , exceeds the threshold, h . The number of responses n , up to time t is then given by

$$n = \frac{e^{-k(h - \epsilon_0 - E)} (1 - e^{-k\theta E t})}{2kE\theta}, \quad t \leq 1/\theta$$

$$= \frac{e^{-k(h - \epsilon_0)}}{2kE\theta} [e^{kE} - 1 + kE(\theta t - 1)] \quad , \quad t \geq 1/\theta. \quad (2.1)$$

In Fig. 2.1, there is shown the agreement of this expression (using the values $\theta = .4$, $kE = 2.73$ and $\exp [-k(h - \epsilon_0)] = 1.9$) with the data for two control animals.

For the experimental situation, it is assumed that the effect of receiving shocks is to produce an inhibition, j , at the rate $\frac{dj}{dt} = \beta - \gamma j$, where w is the number of "wrong" responses (i.e. shocks received). The net excitation is then $\epsilon(t) - j$. After introducing some approximations and simplifications, there results for the number of shocks, w , the expression

$$w = \frac{k}{\beta} \log (1 + k\beta n), \quad t \leq 1/\theta$$

$$= \frac{1}{\beta} e^{-k(h - \epsilon_0 + \beta/\gamma)} \left[t - w\left(\frac{1}{\theta}\right) \right] + w\left(\frac{1}{\theta}\right), \quad t \geq 1/\theta \quad (2.2)$$

Since ψ has been taken equal to $1/w$ ($1/6$), this involves only one additional parameter, k_2 . Using $k_2 = 0.39$, and the values for the other parameters given above, which were estimated for different animals in control runs, gives the curve plotted in Fig. 2.2. This figure also shows the data for a typical experimental animal obtained by Brady and Marmasse.

The constant rate of shocks after $t = 1/w$ given by (2.2) is due to one of the approximations used. Without this, the same model would give a response curve in which the rate decreased asymptotically to a constant value.

In the above it is assumed that conditioning acts only to reduce the tendency of the animal to go to the wired region. There could also be conditioning which results in reduction of general activity. This could be checked by monitoring the activity more carefully. This modification of the experiments has been suggested to Dr. Marmasse.

The model can be made more realistic by including also the fact that there can be a decay of inhibition simply due to forgetting with the passage of time. Also experiments in which shocks are delivered only a fraction of the times that the response is made, can be treated by a slight modification of this model.

Consider next an application to the data by Solomon and Wynne (Psychol. Monog. 1953, 67 No. 4). In this situation dogs were shocked if they did not escape over a barrier within 10 seconds of a warning stimulus. Latencies were recorded. An escape before 10 seconds is referred to as an avoidance while escape after 10 seconds is referred

to as a shock trial.

Let \mathcal{E} denote the excitation level corresponding to the escape response. In the absence of training let it have a value $\mathcal{E}_0 + \mathcal{F}$ where \mathcal{F} represents a random fluctuation. As a result of punishment following the presentation of the conditioned stimulus S_c , there will be a conditioned excitement which will alter \mathcal{F} , so that \mathcal{F} will no longer be, for example, on the average zero, but have some mean value \mathcal{Y} plus the original random component \mathcal{F} .

Let \mathcal{Y} be proportional to the number of shocks, w , for example, so that $\mathcal{Y} = \rho w$. We use this hypothesis rather than postulating that the standard deviation of the fluctuations increase, since this is simpler. In addition, as a result of escape, there may also occur a conditioning between S_c and the escape response.

Let this latter be proportional to the number, c , of escapes, $b c$. Since punishment may lead to the correct response there may also be conditioning of the stimulus S_c following errors, but the effect may not be the same in magnitude, and hence we introduce a coefficient instead of b . Hence we have \mathcal{E} given by the following expression:

$$\mathcal{E} = \mathcal{E}_0 + b c + \rho w + \mathcal{F}. \quad (2.3)$$

If h' is the threshold for the escape response and if we again approximate the distribution function for the random fluctuations by an absolute value exponential, then the probability of a correct response P_0 will be given as a function of w and c by integrating the

fluctuations over the proper range. If we assume that $dc/dn = P_c$, then the expression for P_c can be solved, since $n = w + c$, and hence we may obtain $c(n)$. On differentiating this result with respect to, n , we obtain P_c as a function of n : ($\bar{\beta} = \beta + \rho$)

$$P_c = \frac{\bar{\beta}}{2 \bar{\beta} e^{-\bar{h}} + (\bar{\beta} - \bar{b})(e^{\bar{\beta} n} - 1)}, \quad P_c \leq \frac{1}{2} \quad (2.4)$$

$$P_c = \frac{\bar{\beta}}{2 \bar{b} e^{-\bar{h}} - (\bar{b} - \bar{\beta}) w_2 + \bar{b} n + (\bar{\beta} - \bar{b})(e^{\bar{\beta}(n - n_2)} - 1)}$$

where w_2 and n_2 are related by $n = c + w$,

$$(\bar{\beta} - \bar{b}) c_2 = \bar{\beta} n_2 - \bar{h}. \quad (2.5)$$

In Figure 2.3, data of Solomon and Wynne are used to illustrate the above model. It can be seen that whether b is zero or not results in only a slight improvement for P_c nearly equal to one. Thus on the basis of this model, the most important quantity is $\bar{\beta}$ and, since probably $\beta \approx b$, it is likely that it is the quantity ρ which is the most essential parameter, ρ being the rate of increase per punishment of the conditioned excitement.

Latency in an Avoidance Learning Situation. In view of the previous result, let us neglect b . Now suppose that the time T following the stimulus S_c until the punishment starts, is divided into M parts. During each time period we may write, in a way similar to that outlined

above, an expression for the probability of avoidance, and from this we may then calculate the probability of avoidance during the interval T . The complement of this probability is, by hypothesis, approximately dv/dn and is a function of w only. On integrating the resulting expression we find

$$\bar{\beta}_n = \int_{\log \bar{P}_0}^{(\log \bar{P}_0) e^{\bar{\beta} w}} y^{-1} e^y dy \quad (2.6)$$

The probability $q(t)$ of escape between t and $t + dt$ is given by

$$q(t) = \frac{1}{T} (-\log \bar{P}_0) e^{-\frac{t}{T} \log \bar{P}_0} \quad (2.7)$$

The mean time of escape \bar{T} can be obtained from

$$\frac{\bar{T}}{T} = 1 - P_0^{-1} - \left[-\log \bar{P}_0 \right]^{-1} \quad (2.8)$$

Since the integral of $q(t)$, $Q(t)$, is given by

$$Q(t) = 1 - e^{-\frac{t}{T} \bar{K} - \log 2 + \bar{\beta} w} \quad (2.9)$$

the median time is given from $Q(t_m) = \frac{1}{2}$ above.

If escape does not occur before time T , then shock or punishment occurs. Assume that shock results in an added excitation v' . Then the above argument can be repeated but now we have v' added. In this way we obtain the latency distribution as a function of trials.

3. Application to Free-Recall Verbal Learning.

In previous work (N. Y. Acad. Sc. 96, 1956) it was pointed out that free-recall verbal learning data of Bruner, Miller and Zimmerman (J. Exp. Psychol., 1955, 49) can be accounted for in terms of a single learning coefficient and initial threshold. Tabled values for the percentage correct were given. However, two substantial errors occur in the results given which were not caught in the proof. For 16 and 20 trials, the calculated values should have been 95 and $97\frac{1}{2}$ respectively. In order to correct this error we give here in Figure 3.1 the calculated curves, together with the data.

The situation involved is that in which a subject, after hearing a list of words, is asked to write down all those he can recall. From such records, the proportion of correct responses is calculated for each trial. If we ignore the effect when the subject hears the word but fails to recall it, then the result is as shown in the figure which also contains the theoretical expression in which the threshold $k h'$ is .57 and the learning parameter ($b k$) is .16.

It can be seen that the agreement between theory and experiment is satisfactory. However, it would be of considerable interest to attempt to determine the coefficient $k \beta$, neglected here, but which arises from learning due to hearing the word but without recall. Similarly the effects of interaction between words needs to be considered. A study of the results from various list lengths in the study referred to above needs to be carried out.

Relations between neural net models and stochastic models for learning.

The theories for psychophysical discrimination and learning originated and developed by Rashevsky and Landahl are based on neural net models, and use as basic variables the excitatory and inhibitory impulses in hypothetical nerve pathways and centers. A different approach to mathematical models for learning has been extensively developed in recent years under the general name of stochastic models for learning. The original proposals for these stochastic models were put forward by Bush, Mosteller and Estes. Actually two models are involved here: one considers the samples drawn from the stimulus space by the subject during the learning trials while the other is more formal and considers the probability of a response at any trial as being given by a linear operator, which depends on the response and outcome of the preceding trial, and operates on the probability of response at the preceding trial. These two approaches have been shown to be mathematically equivalent, so that it is correct to consider the stochastic theories of learning together.

In the neural net theory, random fluctuations in the response enter through the consideration of fluctuations in the thresholds and excitations in the nerve centers involved. Thus the question arises of the relation between the neural net theory and the stochastic theories, especially since the stochastic theories involve no statement or hypothesis about any interval events in the subject. This question was considered by H. Martinez in his thesis "Studies in Stochastic Learning Theory" (Committee on Mathematical Biology, 1963). Martinez

showed that, under certain assumptions about the probability distributions in the neural net theory and about the stimulus samples drawn by the subject, it is possible to establish a formal equivalence between the two theories for one particular case.

It is proposed to investigate the relations between these theories further. It is believed that the restrictive assumptions about the distributions and stimulus samples may not be necessary, and that it should be possible to establish the equivalence of the theories under more general conditions. In any case it would be useful to know what assumptions are necessary to prove the two theories equivalent, and under what conditions, if any, they might be expected to give essentially different results. Also the relations between the two theories should be examined in more detail; i.e., when parameters are fitted to the same set of experimental data by both theories, what is the relation between the parameters. What light do these relations throw on the significance of the parameters in both the theories? It is also possible that consideration of the relations between the theories may suggest experiments for discriminating between them, and to help in the interpretation of the parameters.

Miscellaneous Results.

The problem of the b-wave of the electroretinogram has been considered recently (Bull. Math. Biophysics, 25, 125). For the amplitude of the a-wave a simpler model can be considered. Work being carried out in the Department of Ophthalmology, University of Chicago, indicates that this model is satisfactory. The model postulates a generator substance which is diffusible and in the nature of a synaptic coenzyme necessary for membrane depolarization to be able to occur.

In a situation in which a signal is to be detected in the presence of noise it can be shown that an averaging device can greatly increase the certainty of detection. In neural net terms an averaging device is not very simple. However, if one simply asks for sum, with a possible loss, then the result is simply achieved. It can be shown that only summation is needed to give the desired result. Furthermore, if the loss is exponential, a simple expression can be obtained for the probability of a correct response as a function of the number of repetitions. While there are data on externally produced noise levels, data in which there is only internal noise have not been found. The question of the "additivity" of internal and external noise is of interest.

The problem of discrimination among several stimuli is of interest. A particular mechanism has been considered in which there always results the choice of only a single stimulus from among a group of stimuli.

While this mechanism has some undesirable specificity, it may be a starting point for further work.

During the period of this report, the following papers, completed in the previous period (Air Force Contract AF 49(638) - 414) have been published. Copies of these papers, which bear acknowledgement to the Air Force have been mailed out to those individuals and organizations of the current distribution list.

"Mathematical Theory of The Central Nervous System" H. D. Landahl. Annals N. Y. Acad. of Sciences, Vol. 96, Art. 4, 1056-1070, 1962.

"A Canonical Form for Neural Nets Without Circles." J. Towser. Bull. Math. Biophysics, 24, 335-343, 1962.

"On Visual Adaptation I - Photochemistry." Harold White. Bull. Math. Biophysics, 24, 351-359, 1962.

"On Visual Adaptation II - The Electroretinogram and the Bipolar Cells." Harold White. In Press.

FIGURE 1.

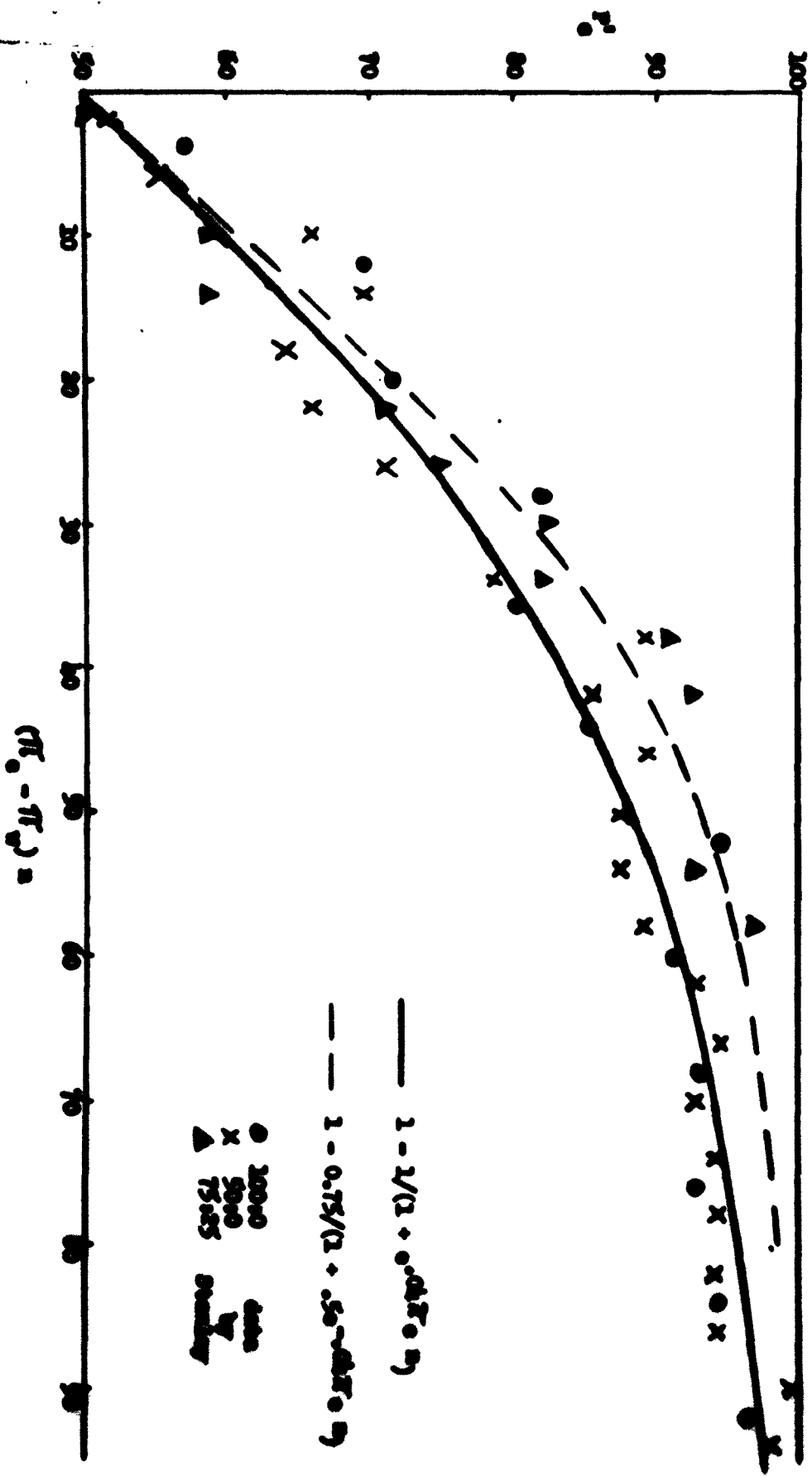
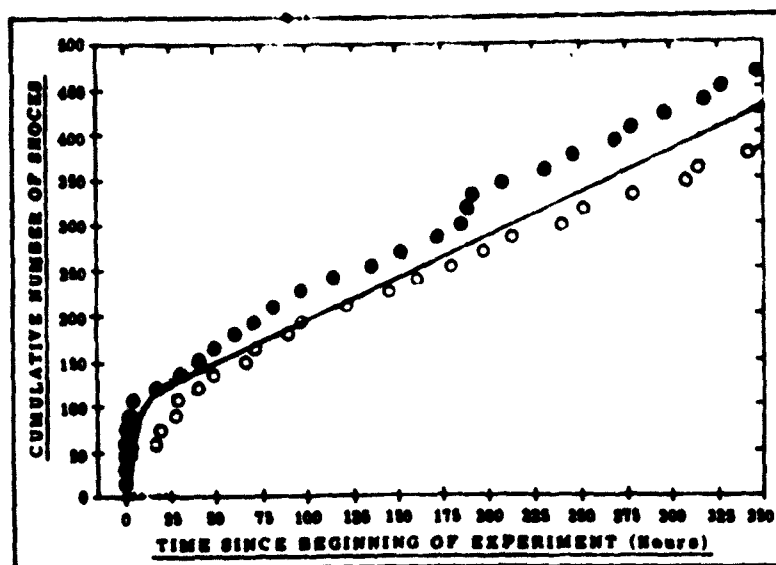
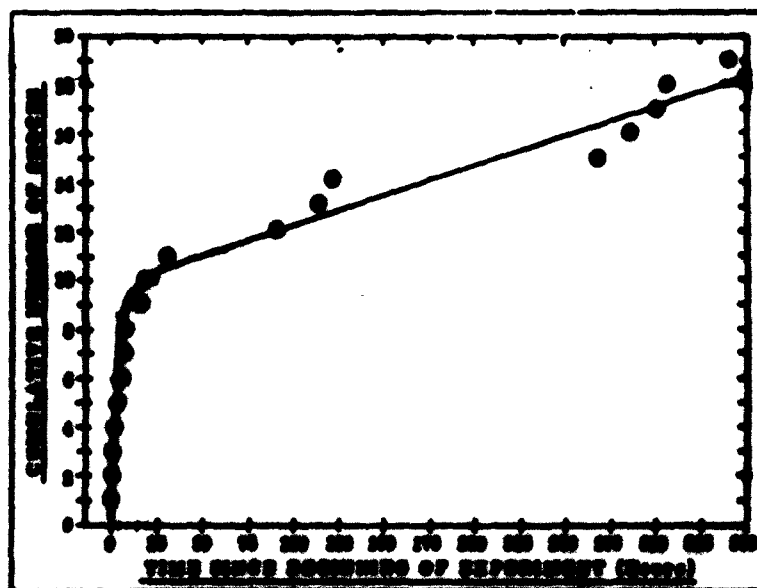


FIGURE 2.1



Data from two control (min-shock) animals (9 and 14). Because of the large number of responses, only cumulative multiples of 15 are plotted.

FIGURE 2.2



Data collected from a typical experimental animal (number 17).

FIGURE 2.3

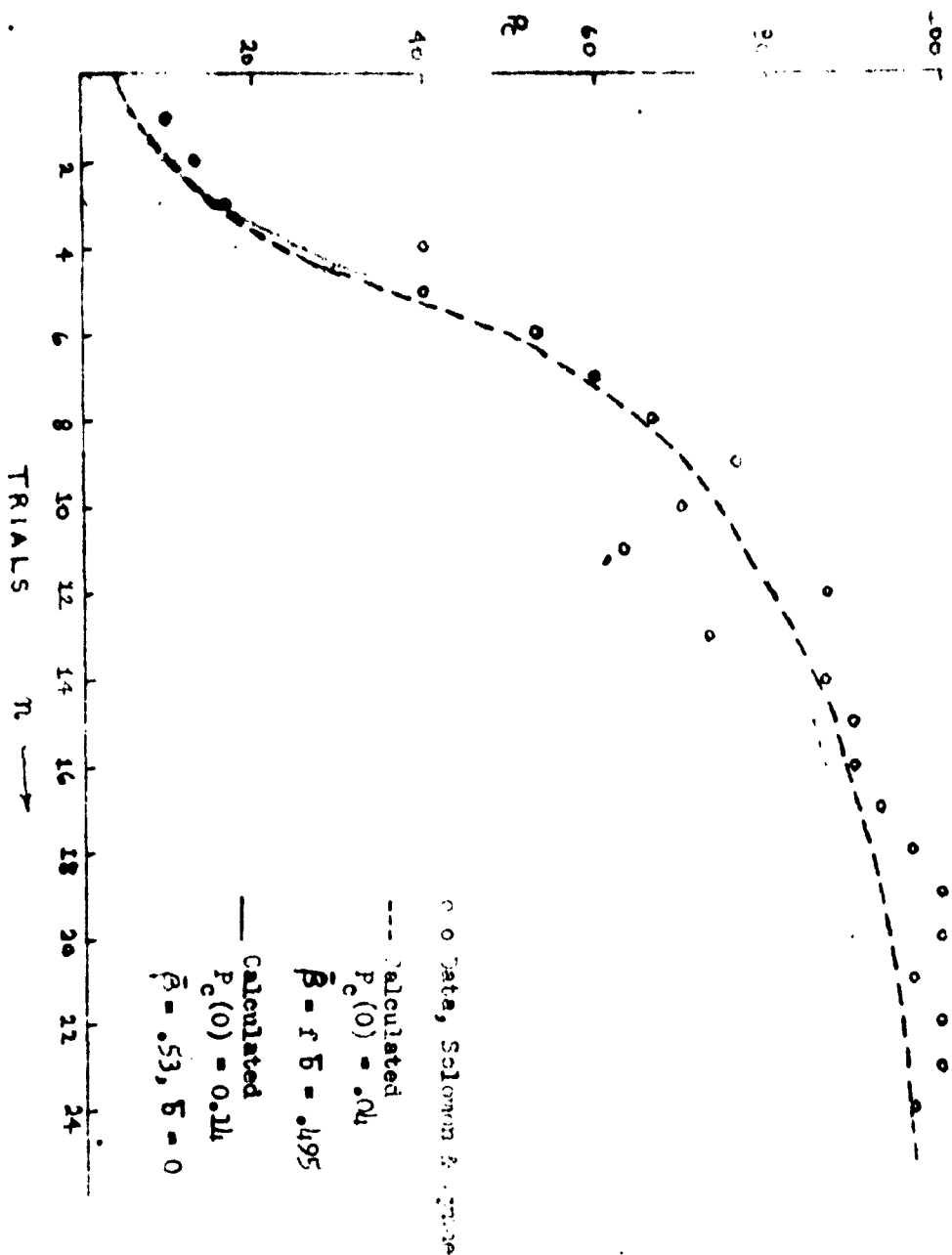
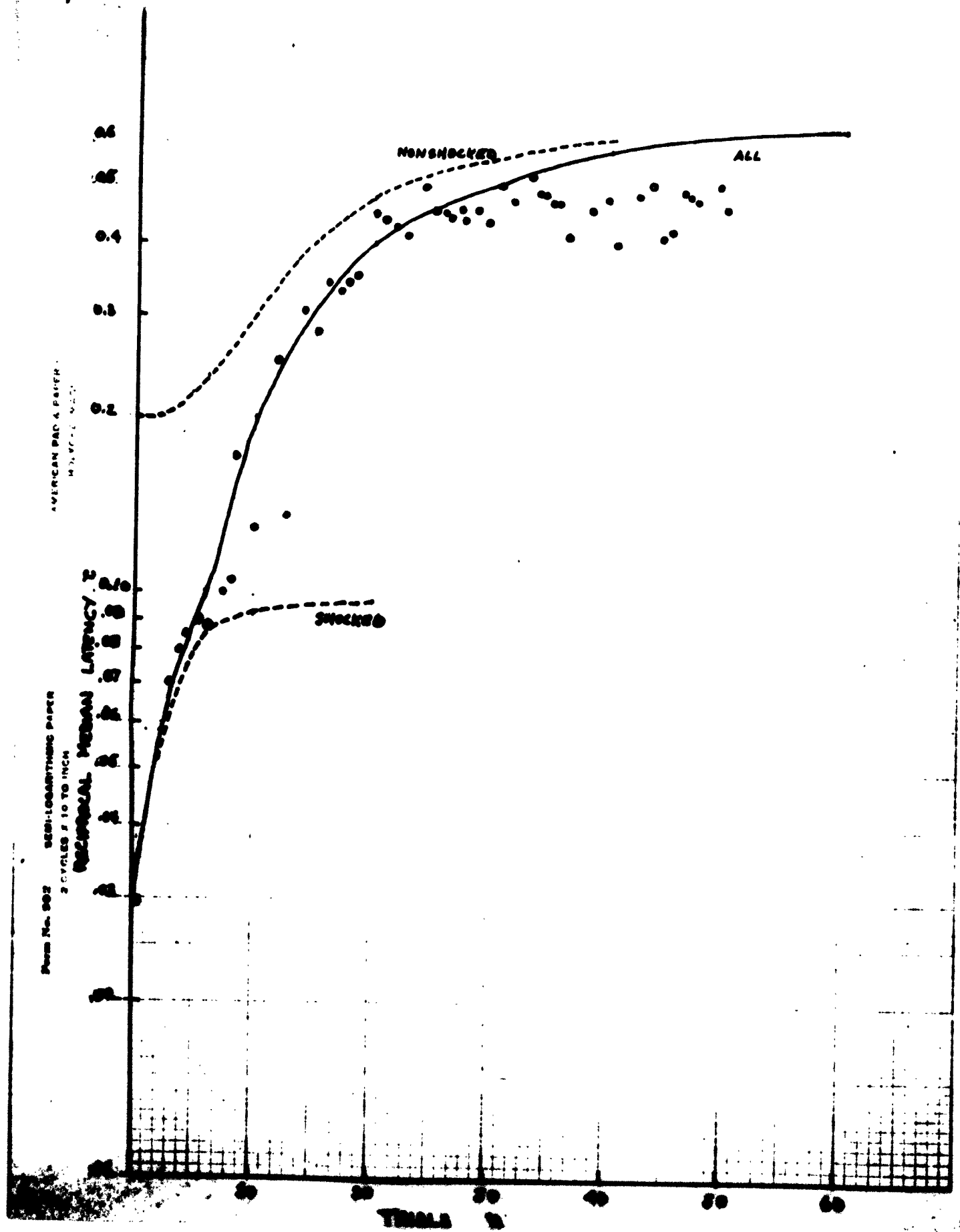
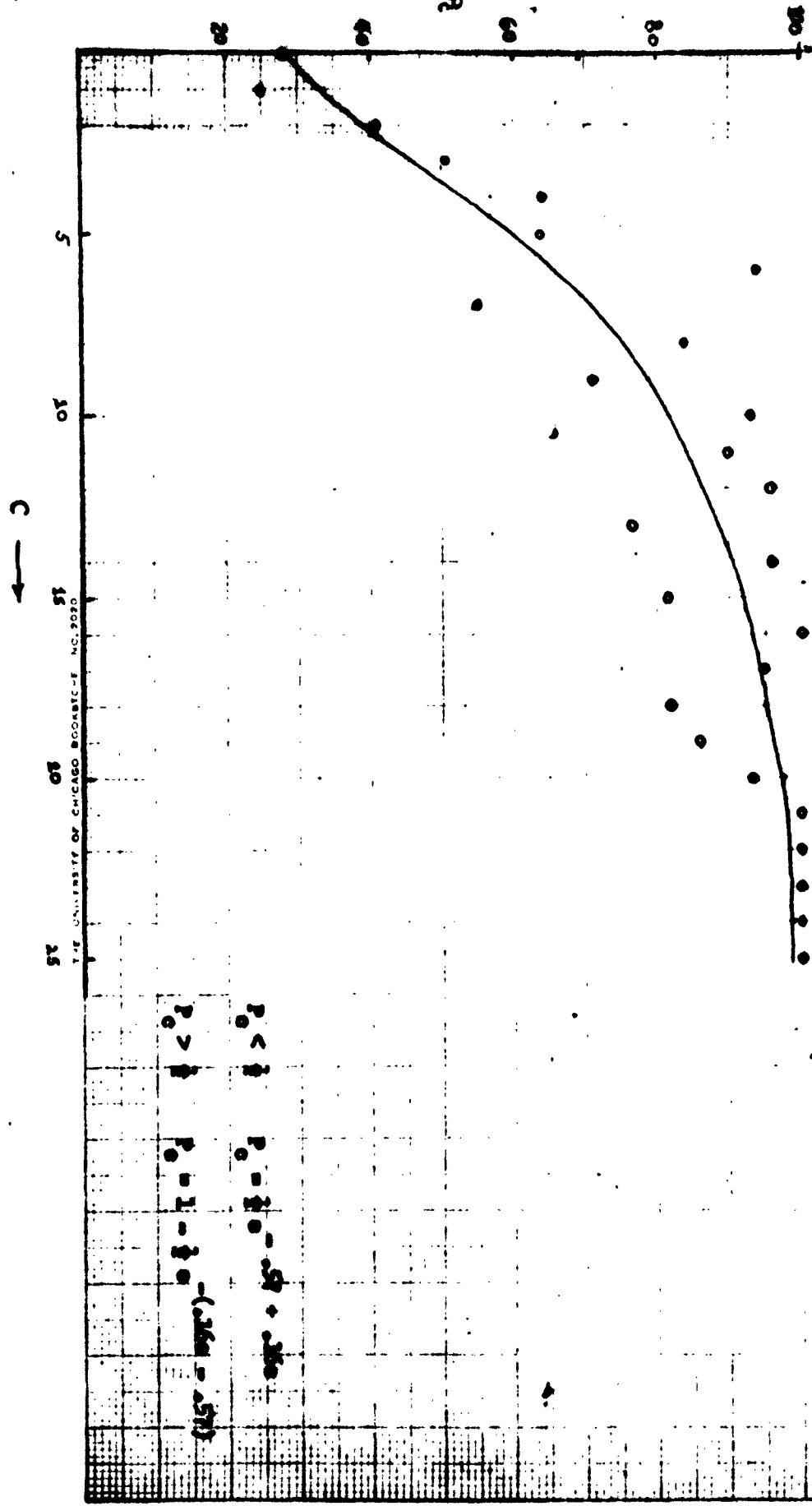


FIGURE 2.4





$$P_0 < \frac{1}{2} \quad P_0 = \frac{1}{2} \quad P_0 = 1 - \frac{1}{2} \quad (1.57) = 0.57$$

THE UNIVERSITY OF CHICAGO BIOLOGICAL MUSEUM